ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 5, pp. 445–456.

Statistical Study of the Quality of the Cycle Merging Algorithm for Solving the Traveling Salesman Problem at Minimum

Yu. F. Leonova[*,a] **and A. V. Panuykov**[*,b]

[*]*South Ural State University, Chelyabinsk, Russia*
e-mail: [a]*yuliya.igosheva@gmail.com,* [b]*paniukovav@susu.ru*

Received March 11, 2024
Revised February 28, 2025
Accepted March 14, 2025

Abstract—The traveling salesman problem is one of the most widely studied problems in the field of combinatorial optimization. However, the investigation of new approaches and the improvement of existing methods remain a significant area of research. This paper presents an analysis of the quality of the cycle merging algorithm used to solve the minimum traveling salesman problem. The results of a computational experiment on five families of problems are provided, and the accuracy and time complexity of the algorithm are analysed. For symmetric instances, a regression model is constructed to describe the dependence of the relative error estimate on the number of vertices. It is shown that a polynomial model provides the best approximation of the obtained data and satisfies key statistical assumptions. The results allow us to evaluate the error growth patterns and to justify the applicability of algorithms to large-scale instances of the traveling salesman problem.

Keywords: traveling salesman problem, heuristic algorithm, asymptotic accuracy of the algorithm, statistical study

DOI: 10.31857/S0005117925050067

1. INTRODUCTION

The traveling salesman problem is a classical combinatorial optimization problem and remains one of the most important and topical problems in the field of applied mathematics and computer science. Its importance is manifested in many practical spheres, such as logistics, transportation, production, route planning and others. The minimum traveling salesman problem consists in finding the Hamiltonian cycle with minimum sum of weights of edges in a complete graph.

In recent decades, a variety of methods for solving the traveling salesman problem has been widely studied. These include exact algorithms, approximations, and heuristic techniques [1, 2]. However, with the advent of new computational technologies and the constant increase in the amount of data, there is a need to develop more efficient and scalable algorithms to solve this problem.

The traveling salesman problem is NP-hard [3]. Therefore, the development of polynomial approximate algorithms is relevant for its solution.

This paper presents an empirical study and statistical analysis of the accuracy and time complexity of a heuristic algorithm for solving the traveling salesman problem — the cycle merging algorithm. Section 2 characterizes the accuracy and running time estimates of popular efficient algorithms for solving the minimum traveling salesman problem. Section 3 gives a brief description of the algorithm under study. Section 4 describes the problem generation and the computational experiment. The results of the computational experiment on random input data are presented in Section 5: data on the accuracy and solution time of the problems are given. Section 6 provides a statistical analysis of the accuracy of the cycle merging algorithm. Section 7 summarizes the main conclusions of the experimental results.

2. STATE OF THE PROBLEM OVERVIEW

Heuristic and approximate algorithms such as nearest neighbor algorithm, greedy algorithm, nearest insertion and others are widely used in practice due to their simplicity and efficiency. They often show good results for small datasets, but may not provide the optimal solution for high dimensional problems.

Table 1 presents priori estimates of the accuracy and time of some heuristic algorithms [4] for the minimum traveling salesman problem. The upper bounds for the algorithms are calculated as the ratio $f(s)/f(s_O)$, where $f(s)$ is the obtained tour length and $f(s_O)$ is the optimal tour length.

Table 1. Upper bound of accuracy estimation and running time of the algorithms

Algorithm	Upper bound	Operation time
Immediate neighbor [5]	$0.5\,[\log_2 N + 1]$	$O(N^2)$
Immediate double-ended neighbor [6]		
Greedy [5]		$O(N^2)\log N$
Nearest addition [6]	$2 - 2/N$	$O(N^2)$
Nearest insert [6, 7]		
Cheapest insert [6, 7]		$O(N^2)\log N$
Farthest insert [6, 7]		$O(N^2)$
Arbitrary insertion [7]		
Inserting the nearest segment		
Dual minimal spatial tree [8]		
Dual minimal spatial tree modified		
Christofides [9]	$3/2 - 1/N$	$O(N^3)$
Mura curve [10]	$\log N$	$O(N \log N)$
Serpinsky curve [11]		
2-Opt [12]	≈ 2	$O(N^2)$

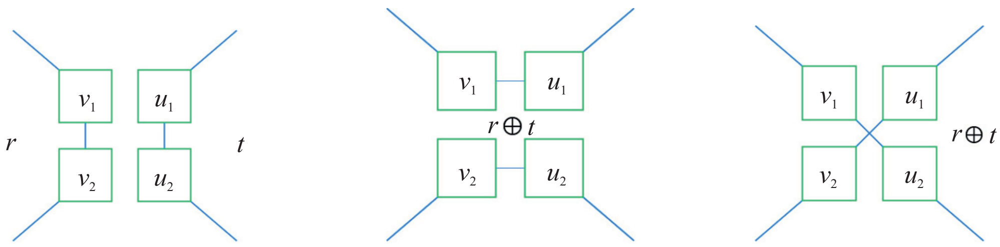
The above algorithms for solving the traveling salesman problem have different advantages and disadvantages depending on the characteristics of the input data and the required accuracy of the result.

The following sections will analyze the accuracy and time complexity of the cycle merging algorithm.

3. THE CYCLE MERGING ALGORITHM FOR SOLVING THE TRAVELING SALESMAN PROBLEM

The authors of the study [13] investigated a heuristic algorithm for solving the traveling salesman problem — Cycle Merging Algorithm (CMA). Here is a brief description of the algorithm.

The first step of the algorithm consists of finding an extreme weight in a given graph of a 2-regular subgraph, i.e. covering this graph with cycles. This construction is commonly referred to as a 2-factor of minimum weight. The problem of finding a 2-factor of minimum weight can be reduced to an assignment problem for which exact polynomial algorithms are known.



Initial cycles $r, t \in C$: $e^r_{\{r\oplus t\}} = \{v_1, v_2\}$, $e^t_{\{r\oplus t\}} = \{u_1, u_2\}$

Merge option 1: $f_{r\oplus t} = \{v_1, v_2\}$, $g_{r\oplus t} = \{u_1, u_2\}$

Merge option 2: $f_{r\oplus t} = \{v_1, v_2\}$, $g_{r\oplus t} = \{u_1, u_2\}$

Fig. 1. Cycle merging options.

In the second phase, the condition of cycle singularity of the obtained solution is checked. If the 2-factor is represented by a single cycle, then this cycle is the solution of the problem. If the 2-factor is represented by several cycles, then different cycles r and t are searched in pairs. In each cycle of the selected pair, one edge is searched.

Let $e^r_{\{r\oplus t\}} = [v1, v2]$ and $e^t_{\{r\oplus t\}} = [u1, u2]$ be chosen, for them two pairs of conjugate edges are found ($f = [v1, u1]$, $g = [v2, u2]$) and ($f = [v1, u2]$, $g = [v2, u1]$), used to connect the cycles (Fig. 1). We search for a set $r, t, e^r_{\{r\oplus t\}}, e^t_{\{r\oplus t\}}, f, g$ of the listed elements such that the connected cycle has minimum weight.

The found pair of cycles is then replaced by the joined cycle with minimum cost. The algorithm terminates when the current 2-factor contains only one cycle.

A theorem on the computational complexity of the algorithm is formulated and proved for the described algorithm.

Theorem 1. The computational complexity of the cycle merging algorithm does not exceed $O(|V|^3)$.

A more detailed description of the cycle merging algorithm and a proof of the computational complexity theorem are given in [13].

For the metric maximal traveling salesman problem, a theoretical accuracy bound was previously established for this algorithm [13], which allows us to rigorously evaluate its quality in this class of problems. However, there are no such theoretical bounds for the minimum traveling salesman problem, so it is of interest to conduct an empirical study of the accuracy of the algorithm.

4. DESCRIPTION OF TEST PROBLEMS

To evaluate the quality of the cycle merging algorithm for the traveling salesman problem at minimum, a computational experiment was conducted on problem sets from the TSPLIB library [14], as well as on randomly generated instances with different characteristics of the cost matrix.

The performance of the algorithm was evaluated on the following families of the traveling salesman problem:

1. Asymmetric instances of the traveling salesman problem from TSPLIB. TSPLIB contains real and artificially generated problems that are widely used in heuristic and exact algorithms. The asymmetric cost of traveling between vertices makes the solution more complex, since the inherent symmetry inherent in many classical methods is broken.

2. Euclidean instances of the traveling salesman problem from TSPLIB ($n \leqslant 3000$). This data type models geometric versions of the problem, where vertices correspond to points in the plane and

the costs of transitions between them are determined by the Euclidean distance. Such problems are commonly found in logistics, navigation, and network planning.

3. Asymmetric cost matrices $W = [w(i,j)]$, where $w(i,j)$ are random numbers from $\{0, 1, 2, \ldots, 10^5\}$. The uniform distribution of values in a fixed interval models random systems with independent transition weights. This allows to test the algorithm under full uncertainty of the structure value.

4. Asymmetric cost matrices where $w(i,j)$ are random numbers from $\{0, 1, 2, \ldots, i \times j\}$. In this family, the range of possible costs increases as the vertex indices increase, which models problems where the communication between vertices becomes more expensive or complex as their numbers increase. This type of matrix allows to analyze the behavior of the algorithm on problems with non-uniform complexity of links.

5. Symmetric cost matrices W where $w(i,j)$ are random numbers from $\{0, 1, 2, \ldots, 10^5\}$ for $i < j$. In contrast to the asymmetric cases, the symmetric structure limits the set of possible solutions, but leads to an increase in the number of short cycles in the original coverage. This makes it difficult to combine cycles into a single route and can negatively affect the accuracy of the heuristic algorithm.

6. Symmetric matrices where $w(i,j)$ are random numbers from $\{0, 1, 2, \ldots, i \times j\}$ for $i < j$. Similar to the asymmetric case, but taking into account the symmetry of the weights. This type of problem models cases where distant vertices are associated with higher cost of transitions. However, the presence of many short cycles in the early stages of the solution makes the problem computationally complex, and may worsen the accuracy of an approximate solution.

7. Sloped plane instances of the problem where the arc weight is defined as

$$w(i,j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} - \max(0, y_i - y_j) + 2\max(0, y_j - y_i),$$

with vertex coordinates generated independently and uniformly on the interval $\{0, 1, 2, \ldots, 10^5\}$. This model simulates a situation in which moving vertically involves additional costs, such as traveling on mountain roads. The difference in the cost of upward and downward movement makes the problem asymmetric and complicates its solution.

For families 3–6, the number of vertices n was varied from 100 to 3000 by increments of 100. All results are the average of 100 trials each.

The experiment uses instances that are generated in both a random and determined way. The number and variety of families used allows us to check the reliability of the algorithm under test.

5. RESULTS OF COMPUTATIONAL EXPERIMENT FOR THE TRAVELING SALESMAN PROBLEM AT MINIMUM

For the first and the second family of problem instances (asymmetric and euclidean problems from TSPLIB), the results are given in [15]. The analysis of accuracy and computational efficiency of the cycle merging algorithm for the other families are considered within the framework of this paper.

5.1. Analysis of Relative Error Estimation

Figure 2 shows the relative error estimates obtained in the course of the numerical experiment for different families of traveling salesman problems. The elements of the cost matrix of these problems were randomly generated according to the generation rules described in Section 4.

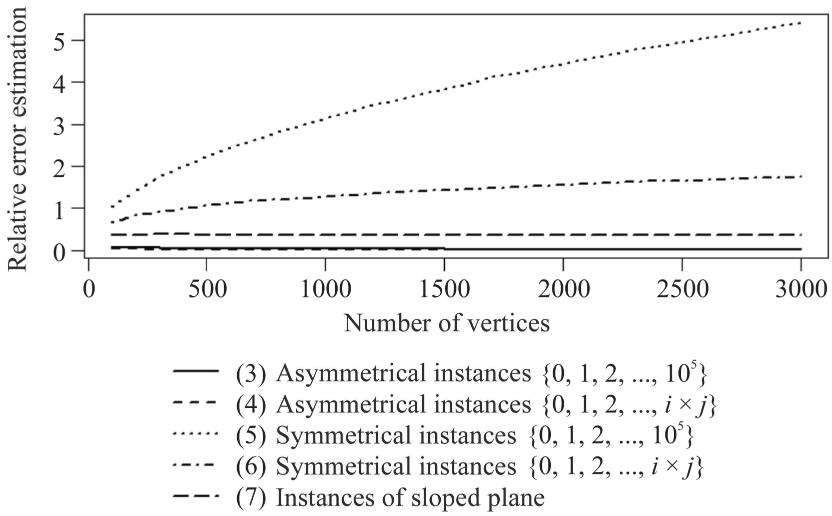


Fig. 2. Relative error estimation of the cycle merging algorithm for different families of the traveling salesman problem.

The analysis of the results shows differences in the behavior of the relative error estimation depending on the family of the considered instances.

For asymmetric instances of the traveling salesman problem with uniformly distributed elements of the weight matrix from the set $\{0, 1, 2, \ldots, 10^5\}$ (family 3) relative error decreases with increasing number of vertices. At $n = 100$ its average value is 9.15%, and at $n = 3000$ it decreases to 1.97%. Analogous tendency is observed for family 4, but the error value in average remains lower, decreasing from 5.12% at $n = 100$ to 1.06% at $n = 3000$.

For symmetric instances of the traveling salesman problem, represented by families 5 and 6, a different behavior is observed. The relative error in both cases increases with increasing number of vertices. Thus, for family 5 it increases from 103% at $n = 100$ to 541% at $n = 3000$. A similar increase is characteristic for the family 6, with the error increasing from 65 to 175% for the same values of n.

For the sloped plane instances (family 7), the estimate of the relative error remains relatively stable over the entire interval of n values, fluctuating between 36.6–37.7%. This result indicates that the accuracy of the algorithm is less dependent on the dimensionality of the problem in this family of instances.

Thus, the analysis shows that:

- for asymmetric instances of the traveling salesman problem, the estimate of the relative error decreases with increasing number of vertices;
- for symmetric instances of the traveling salesman problem, the opposite trend is observed — the error increases with increasing problem dimensionality;
- the family of sloped plane demonstrates stable values of the relative error, insignificantly changing with increasing number of vertices.

5.2. Analyzing the Solution Time

Figure 3 presents the average running time of the cycle merging algorithm at solving one problem for different families of instances of the traveling salesman problem.

The analysis of time characteristics shows different behavior of time growth of the solution depending on the family of considered instances.

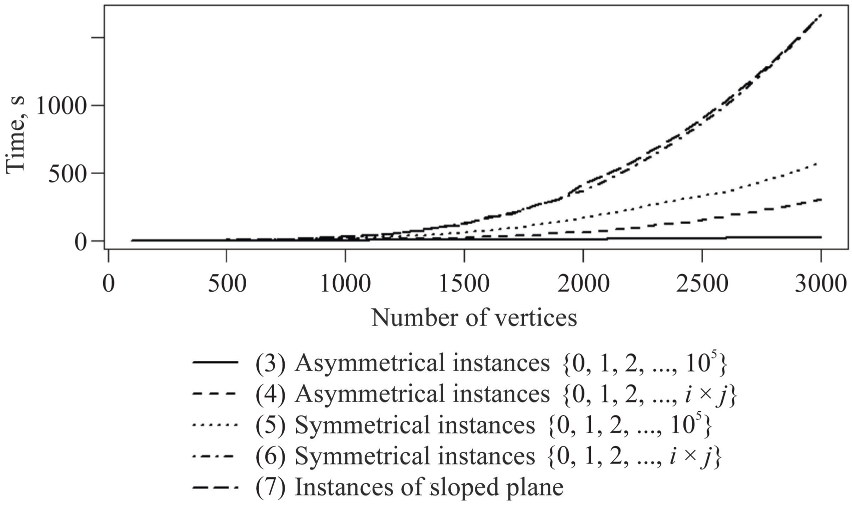


Fig. 3. Solution time per instance of the traveling salesman problem using the cycle merging algorithm for various problem families.

For asymmetric instances of the traveling salesman problem with uniformly distributed elements of the weight matrix from the set $\{0, 1, 2, \ldots, 10^5\}$ (family 3) solution time increases from 0.105 s at $n = 100$ to 259.88 s at $n = 3000$. For family 4 ($W = [w(i, j)]$, where $w(i, j)$ are uniformly distributed random numbers of $\{0, 1, 2, \ldots, i \times j\}$) the solution time grows faster, starting from 0.14 s at $n = 100$ and reaching 3057.04 s at $n = 3000$. This indicates the significant influence of structure of the cost matrix on the computational complexity of the algorithm.

For symmetric instances of the traveling salesman problem, a faster growth of the solution time is observed. In particular, for family 5 ($W = [w(i, j)]$, where $w(i, j)$ are uniformly distributed random numbers from $\{0, 1, 2, \ldots, 10^5\}$) it increases from 0.114 s when $n = 100$ to 5760.3 s when $n = 3000$. For family 6 ($W = [w(i, j)]$, where $w(i, j)$ are uniformly distributed random numbers from $\{0, 1, 2, \ldots, i \times j\}$) this growth of is even more pronounced: from 0.27 s to 16 669.63 s over the same interval of n values. Thus, the specificity of the structure of weights influences not only the accuracy, but also the computational complexity of the algorithm.

For the instances of sloped plane (family 7), a significant increase in the solution time is also observed. At $n = 100$ the average time is 0.279 s, and at $n = 3000$ it reaches 16 694.93 s, which is similar to the indicators for the sixth family of problems and significantly exceeds the data for all other families of problems. This indicates the complexity of processing such instances by the cycle merging algorithm.

In general, the analysis of temporal characteristics shows that:

- asymmetric instances of the traveling salesman problem are solved faster than symmetric ones, but the structure of the cost matrix significantly affects the growth of solution time;
- for symmetric instances there is an accelerated growth of the running time of the algorithm, especially for family 6;
- instances of sloped plane show the highest values of the solution time, which may be related to the peculiarities of the geometric structure of the problems.

The complexity of the symmetric non-Euclidean traveling salesman problem for algorithms starting from the assignment problem is due to the low accuracy of the lower bound obtained using this method. In particular, for the vast majority of symmetric instances, the assignment problem leads to a cycle coverage containing a large number of short cycles (Table 2). This significantly complicates the process of combining them into a single route, increasing both the computational cost and the resulting relative error.

Table 2. Ratio of the number of initial cycles to the number of problem vertices

Number of vertices	Asymmetrical instances	Symmetrical instances
100	0.0382	0.4920
200	0.0225	0.4768
300	0.0162	0.4771
400	0.0139	0.4892
500	0.0118	0.4848
600	0.0101	0.4867
700	0.0086	0.4896
800	0.0076	0.4902
900	0.0071	0.4890
1000	0.0070	0.4878

As can be seen from the table, for asymmetric instances the ratio of the number of initial cycles to the number of vertices decreases significantly with increasing dimensionality of problem, decreasing from 0.0382 (100 vertices) to 0.0070 (1000 vertices). This indicates a tendency to form longer cycles, which contributes to a more efficient connection of routes at subsequent stages of the algorithm.

In contrast, for symmetric instances this ratio remains practically unchanged, fluctuating in the range 0.4768–0.4920. This indicates that during the formation of the 2-factor for this family of problems, a lot of short subcycles are formed, which makes it difficult to combine them into the final route.

The growth of the solution time for symmetric instances is explained by the fact that the process of connecting a large number of short cycles requires a significantly larger number of iterations, which leads to a complication of the algorithm structure and an increase of its computational complexity. As shown in Fig. 3, the running time of the algorithm for connecting cycles for symmetric instances significantly exceeds the similar values for asymmetric problems.

Thus, in order to improve the accuracy and reduce the time to solve the symmetric non-Euclidean traveling salesman problem, further study of specialized heuristics and approximate algorithms adapted to work with this family of problems is required.

6. STATISTICAL ANALYSIS OF THE ACCURACY OF THE CYCLE MERGING ALGORITHM FOR SYMMETRIC TRAVELING SALESMAN PROBLEM

As part of the investigation of the accuracy of the cycle merging algorithm for the symmetric traveling salesman problem at minimum, a statistical analysis of empirical data obtained for instances of the problem with a weight matrix whose elements are generated uniformly from the set $\{0, 1, 2, \ldots, i \times j\}$ (the fifth family of problem instances from Section 4) was carried. Since there are no theoretical bounds on the accuracy of the algorithm for the symmetric minimum comparison problem, we used regression analysis methods to predict the dependence of the relative error on the number of vertices.

Three possible regression models were considered: steppe, second-order polynomial, and logarithmic. These models were chosen on the basis of observations on the nature of the relative error growth with increasing number of vertices.

6.1. *Selection of Regression Model*

Exploratory analysis is a necessary step to test the assumptions [16] associated with regression models and to select the most appropriate model to approximate the data. As part of this analysis phase, visual inspections were performed and statistical characteristics were calculated for each model to compare their performance and accuracy.

Visual methods such as scatter dot plot and dot plot matrix were used to analyze the relationship between the number of vertices and relative error. Histograms and tests for normality of the distribution of residuals were also used. Statistical tests, such as the coefficient of determination R^2, adjusted R^2, and the Durbin–Watson and Breusch–Pagan tests, were used to assess the model's fit to the statistical assumptions.

6.1.1. Power regression. The power model is of the form

$$y = ax^b. \tag{1}$$

After logarithmization, the equation takes a linear form:

$$\log y = \log a + b \cdot \log x. \tag{2}$$

The following results were obtained for this model:

- The coefficient of determination $R^2 = 0.9676$, adjusted $R^2 = 0.9665$, which indicates high quality of data approximation.
- The significance of the model coefficients is confirmed by the p-values, which are less than $2 \cdot 10^{-16}$.
- The Durbin–Watson test showed no significant autocorrelation of the residuals ($DW = 1.9527$, $p = 0.3711$).
- The Breusch–Pagan test revealed the presence of heteroscedasticity ($p = 1.24 \cdot 10^{-6}$), indicating instability in the variance of the residuals.
- The global test for compliance with the linear regression assumptions showed violation of the assumptions ($p = 2.81 \cdot 10^{-7}$).

Thus, although the power model has a high coefficient of determination and provides a good approximation, the identified problems with heteroscedasticity and violation of linear assumptions indicate its limitations.

6.1.2. Polynomial regression. The second-order polynomial model is

$$\log y = a + b \cdot \log x + c \cdot \log^2 x. \tag{3}$$

The statistical characteristics of this model are:

- $R^2 = 0.9999$, adjusted $R^2 = 0.9999$, which confirms the almost perfect fit of the model data.
- All model coefficients are statistically significant ($p < 0.001$).
- The Durbin–Watson test showed no autocorrelation ($DW = 2.2082$, $p = 0.5771$).
- The Breusch–Pagan test showed no heteroscedasticity ($p = 0.1035$).
- The global test showed small deviations from the assumptions ($p = 0.002$).

The second-order polynomial model showed a significantly better fit to the data compared to the power model, eliminating heteroskedasticity problems. However, the test for global fit to the linearity assumptions indicates possible deviations, which nevertheless do not have a significant impact on the accuracy of the model.

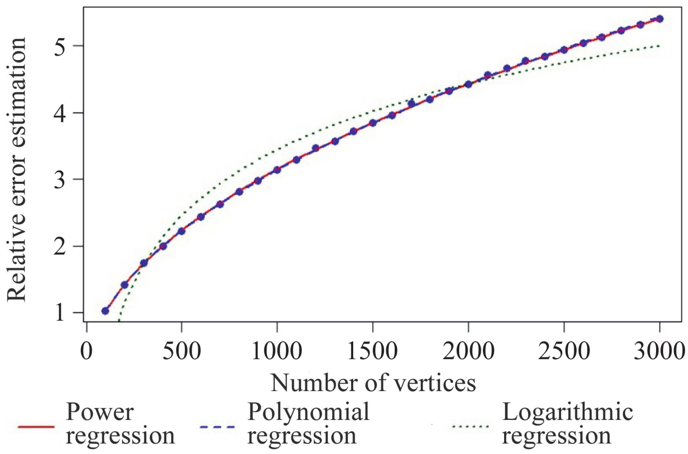


Fig. 4. Comparison of regression models for symmetric instances from the set $\{0, 1, 2, \ldots, 10^5\}$.

6.1.3. Logarithmic regression. The logarithmic model is described by equation

$$y = a + b \cdot \ln x. \tag{4}$$

Statistical results:

- R^2 is slightly lower than that of the polynomial model, but remains high.
- The residuals meet the requirements of normality and homoscedasticity.
- The global test confirms satisfactory compliance with the prerequisites.

The logarithmic model, as well as the polynomial model, showed good correspondence with the data and regression assumptions with easily interpretable coefficients.

6.1.4. Model selection. Figure 4 shows plots of empirical and model values for different personal regression models.

The analysis showed that the power model has a high coefficient of determination, but violates the regression assumptions. The second-order polynomial model significantly improves the quality of approximation by eliminating heteroscedasticity. The third-order polynomial model improves the quality of predictions, but may lead to overcomplication. The logarithmic model describes the slowing growth well, but has limitations in extrapolation.

According to the test results and interpretation of the data, the second-order polynomial model and logarithmic regression are the most fitting. The polynomial model provided the best fit to the data, eliminating the major problems of power regression. The logarithmic model describes slowing growth well, but when extrapolated to values beyond the training set, it can produce bias because the logarithmic relationship assumes infinite growth at large x, which does not always fit real data. In the further analysis we will use the polynomial model of the second order as the most accurate and satisfying the key statistical requirements (Table 3).

Table 3. Global verification of polynomial regression assumptions

Test	Value	p-value
Global Stat	16.843	0.0021 (not satisfied)
Skewness	3.189	0.0742 (acceptable)
Kurtosis	2.321	0.1277 (acceptable)
Link Function	9.898	0.0017 (not satisfied)
Heteroskedasticity	1.436	0.2308 (acceptable)

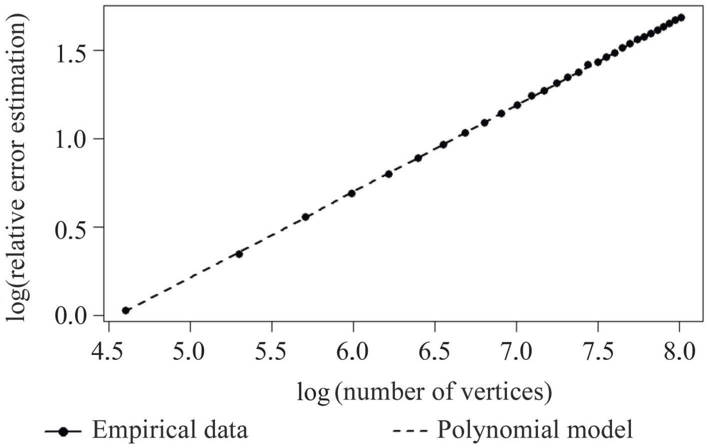


Fig. 5. Empirical and modeled values of logarithmically transformed data for polynomial regression on symmetric instances from the set $\{0, 1, 2, \ldots, 10^5\}$.

The polynomial model on a logarithmic scale shows a good fit to the data, but small deviations from linearity may indicate its limitations in extrapolation. In general, the results support their applicability, although a possible bias should be considered when forecasting large values.

6.2. Evaluation and Interpretation of the Model

The constructed polynomial regression model describes the dependence of the average relative error on the number of vertices. The inclusion of a quadratic term made it possible to take into account the nonlinear nature of the error growth. The coefficient $b_1 = 0.4465$ at $\log n$ indicates an accelerating growth of the error with increasing dimensionality of the problem, and the positive value $b_2 = 0.0035$ at $(\log n)^2$ confirms a small acceleration of growth at small values of n. However, this effect is smoothed out at large problem dimensions, which indicates a tendency to stabilization of error.

Key statistical measures of the model:

- $R^2 = 0.9999$ (almost complete explanation of variance).
- F-statistic: 129 100 ($p < 2 \cdot 10^{-16}$), which confirms the significance of the model and high degree of fit to the data.
- Confidence intervals show that all model coefficients are statistically significant ($p < 0.01$).

Figure 5 presents a graphical comparison of empirical data and predicted values of the model.

Analysis of the global regression assumption test showed a violation of the link function ($p = 0.0017$). However, the absence of heteroscedasticity and autocorrelation of residuals allows us to consider the polynomial model suitable for describing the relationship, despite small deviations from the assumptions of linearity.

6.3. Prediction for Problems of High Dimensionality

On the basis of the constructed polynomial model, we extrapolated to predict the relative error for problems of up to 10 000 vertices (Fig. 6). The results show that at large n the tendency to the error growth remains, but the effect of acceleration due to the quadratic term becomes less pronounced. This confirms that the relative error tends to stabilize and its increase slows down.

Statistical analysis demonstrates that the proposed polynomial model describes well the dependence of the relative error on the dimensionality of the problem and allows us to make predictions

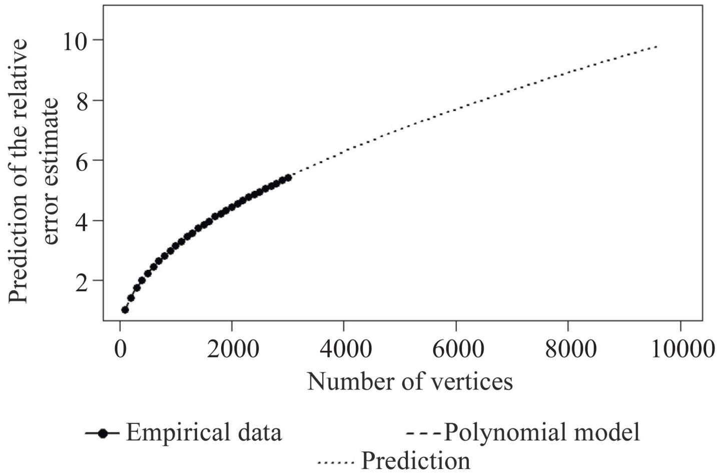


Fig. 6. Prediction of the relative error estimate for large-scale symmetric instances from the set $\{0, 1, 2, \ldots, 10^5\}$.

for large n. Although the model retains high approximation accuracy, possible deviations in extrapolation indicate the need for further investigation of error behavior for problems of extremely large dimensions. Nevertheless, the obtained results confirm that the cycle merging algorithm demonstrates a stable character of change of error and can be applied to the solution of the symmetric salesman problem in the case of high dimensionality.

7. CONCLUSION

The article presents the results of an empirical study of the accuracy and time complexity of the cycle merging algorithm, as well as a statistical analysis of the accuracy of the algorithm for one of the considered families of the symmetric traveling salesman problem on the minimum.

The empirical analysis was performed on seven families of problems, of which the article presents results for five families of randomly generated instances.

The study shows that the solution quality and computational complexity of the algorithm depend on the structure of the cost matrix. For asymmetric instances the relative error estimate decreases to 1–2% at 3000 vertices, whereas for symmetric instances and sloped plane instances there is a gradual increase in relative error. The solution time for all families increases polynomially, with the solution time for symmetric and sloped plane instances being significantly higher than for the asymmetric instances.

The analysis shows that the relative error of the cycle merging algorithm for the symmetric traveling salesman problem demonstrates polynomial growth with increasing dimensionality of the problem. This is confirmed by statistical processing of experimental data, where the polynomial regression model provides the best fit to empirical observations.

In spite of the error growth, the proposed algorithm retains acceptable accuracy and polynomial running time, which makes it a promising tool for solving instances of the traveling salesman problem of large dimensionality. To further improve the quality of solutions, it is advisable to investigate hybrid approaches combining cycle merging algorithm with additional heuristics and local improvements.

REFERENCES

1. Zhang, C. and Sun, P., Heuristic methods for solving the traveling salesman problem (TSP): a comparative study, *Proc. IEEE 34th Annu. Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 2023, pp. 1–6. https://doi.org/10.1109/PIMRC56721.2023.10293957

2. Toaza, B. and Esztergár-Kiss, D., A review of metaheuristic algorithms for solving TSP-based scheduling optimization problems, *Appl. Soft Comput.*, 2023, vol. 148, art. no. 110908. https://doi.org/10.1016/j.asoc.2023.110908

3. Reingold, E.M., Nievergelt, J., and Deo, N., *Combinatorial Algorithms: Theory and Practice. Englewood Cliffs: Prentice Hall*, 1977. https://doi.org/10.2307/2987917

4. Avdoshin, S.M. and Beresneva, E.N., Metric traveling salesman problem: experimental study of Pareto-optimal algorithms, *Proc. of ISP RAS*, 2017, vol. 29, no. 4, pp. 123–138. https://doi.org/10.15514/ISPRAS-2017-29(4)-8

5. Flood, M.M., The traveling-salesman problem, *Oper. Res.*, 1956, vol. 4, no. 1, pp. 61–75. https://doi.org/10.1287/opre.4.1.61

6. Rosenkrantz, D.J., Stearns, R.E., and Lewis II, P.M., An analysis of several heuristics for the traveling salesman problem, *SIAM J. Comput.*, 1977, vol. 6, no. 3, pp. 563–581. https://doi.org/10.1137/0206041

7. Hahsler, M. and Hornik, K., TSP — Infrastructure for the traveling salesperson problem, *J. Statist. Softw.*, 2007, vol. 23, no. 2. https://doi.org/10.18637/jss.v023.i02

8. Kay, E. and Christofides, N., Graph theory: an algorithmic approach, *Oper. Res. Q.*, 1976, vol. 27, no. 4, pp. 1027. https://doi.org/10.2307/3009122

9. Christofides, N., Worst-case analysis of a new heuristic for the traveling salesman problem, Pittsburgh, PA: Graduate School of Industrial Administration, Carnegie Mellon Univ., 1976. https://doi.org/10.1007/s43069-021-00101-z

10. Buchin, K., Space-filling curves, *Organizing Point Sets: Space-Filling Curves, Delaunay Tessellations of Random Point Sets, and Flow Complexes*, Berlin: Free Univ. of Berlin, 2007, pp. 5–30.

11. Bartholdi, J.J., Platzman, L.K., Collins, R.L., and Warden, W.H., A minimal technology routing system for meals on wheels, *Interfaces*, 1983, vol. 13, no. 3, pp. 1–8. https://doi.org/10.1287/inte.13.3.1

12. Aarts, E.H.L. and Lenstra, *J.K., Local Search in Combinatorial Optimization. Princeton*, NJ: Princeton Univ. Press, 2003. https://doi.org/10.1515/9780691187563

13. Panyukov, A.V. and Leonova, Y.F., Cycle merging algorithm for the metric traveling salesman problem for maximum, *Bull. South Ural State Univ., Ser. Math. Model. Program. Comput. Software*, 2021, vol. 10, no. 4, pp. 26–36. https://doi.org/10.14529/cmse210402

14. Reinelt, G., TSPLIB—A traveling salesman problem library, *INFORMS J. Comput.*, 1991, vol. 3, no. 4, pp. 376–384. https://doi.org/10.1287/ijoc.3.4.376

15. Leonova, Y.F., Analysis of the cycle merging algorithm quality on TSPLIB instances, *Actual Problems of Applied Mathematics, Computer Science and Mechanics. Proc. Int. Sci. Conf.*, Voronezh, Dec. 13–15, 2021, Voronezh: VSU, 2022, pp. 573–579. (In Russian)

16. Heins, J., Bossek, J., Pohl, J., Seiler, M., Trautmann, H., and Kerschke, P., A study on the effects of normalized TSP features for automated algorithm selection, *Theor. Comput. Sci.*, 2023, vol. 940, pp. 123–145. https://doi.org/10.1016/j.tcs.2023.01.045

This paper was recommended for publication by A.A. Lazarev, a member of the Editorial Board